

EV316937787

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

**System and Method for
Associating Identifiers with Data**

Inventors:

Jeremy N. Sokolic

Balraj Suneja

Amitava Parial

Sarabjeet Singh

Gautam Sinha

Sanjeev Dheer

ATTORNEY'S DOCKET NO. CE1-012US

1 **RELATED APPLICATIONS**

2 This application is a continuation-in-part of co-pending application Ser. No.
3 10/040,314, filed January 3, 2002, entitled "Method and Apparatus for Retrieving
4 and Processing Data", and incorporated herein by reference.

5
6 **TECHNICAL FIELD**

7 The present invention relates to associating identifiers with data, such as
8 financial data.

9
10 **BACKGROUND**

11 Individuals, businesses, and other organizations typically maintain one or
12 more financial accounts at one or more financial institutions. Financial institutions
13 include, for example, investment institutions, life insurance vendors, banks,
14 savings and loans, credit unions, mortgage companies, lending companies, and
15 stock brokers. Financial accounts may include asset accounts (such as brokerage
16 accounts, investment accounts, 401k accounts, other retirement accounts, mutual
17 fund accounts, life insurance and annuity accounts, bank savings accounts,
18 checking accounts, and certificates of deposit (CDs)) and liability accounts (such
19 as credit card accounts, mortgage accounts, home equity loans, overdraft
20 protection, and other types of loans). Liability accounts may also be referred to as
21 "debt accounts".

22 Many financial institutions allow customers to access information regarding
23 their accounts via the Internet or other remote connection mechanism (often
24 referred to as "online banking"). Typically, the customer navigates, using a web
25 browser application, to a web site maintained by the financial institution. The web

1 site allows the customer to login by entering a user identification and an associated
2 password. If the financial institution accepts the user identification and password,
3 the customer is permitted to access information (e.g., account holdings and
4 account balances) regarding the financial accounts maintained at that financial
5 institution.

6 Similarly, other organizations and institutions allow customer access to
7 other types of accounts, such as email accounts, award (or reward) accounts,
8 online bill payment accounts, etc. A user may navigate a web site or other
9 information source to receive status information regarding one or more of their
10 accounts.

11 Account information (such as information regarding publicly traded
12 financial securities held as investment positions and account transactions)
13 associated with different financial institutions may have different identifiers
14 associated with the account information. Data collected regarding investment
15 securities, such as data gathered from different web-based online financial
16 accounts, often lacks a standard unique identifier. For example, some data sources
17 provide a ticker symbol, but the ticker symbol is neither unique nor consistent
18 from one data source to another. For some securities there are no ticker symbols.
19 For example, one data source (e.g., a brokerage firm) may list a security's ticker as
20 "ACME.A" while another data source uses a different ticker ("ACME_A") for the
21 same security. Other data sources may use "ACME'A" or "ACME A" for this
22 same security. Further, the name assigned to the security may vary from one data
23 source to another. For example, for the above security, different data sources may
24 name the security "ACME SYSTEMS INC CL A", "Acme Systems A", or
25

1 “ACME SYSTEMS class A” – all identifying the same class A common stock
2 associated with Acme Systems Inc.

3 In other situations, a data source may not provide any ticker symbol or
4 other identifier for a particular security. As mentioned above, the name assigned
5 to the same security may vary from one data source to another. These
6 inconsistencies lead to difficulties in properly identifying and handling
7 information regarding securities when the information is collected from multiple
8 sources.

9
10
11
12
13
14

1 **BRIEF DESCRIPTION OF THE DRAWINGS**

2 Fig. 1 illustrates an example network environment in which various servers,
3 computing devices, and a financial analysis system exchange data across a
4 network, such as the Internet.

5 Fig. 2 is a block diagram showing example components and modules of a
6 financial analysis system.

7 Figs. 3A and 3B illustrate a flow diagram of a procedure for retrieving data
8 and associating identifiers with the retrieved data.

9 Fig. 4 is a flow diagram illustrating a procedure for retrieving data and
10 associating asset identifiers with the retrieved data based on various rules.

11 Fig. 5 is a flow diagram illustrating a procedure 500 for applying various
12 rules or search patterns to determine an identifier associated with a data element.

13 Fig. 6 illustrates an example set of rules used to associate data elements
14 with identifiers.

15 Fig. 7 is a block diagram showing pertinent components of a computer in
16 accordance with the invention.

1 **DETAILED DESCRIPTION**

2 The systems and methods described herein are capable of retrieving and
3 handling data from one or more data sources, such as financial institutions. In
4 particular, these systems and methods are capable of assigning a common set of
5 identifiers to aggregated data using rules that contain, for example, information
6 regarding financial securities, financial institutions, financial institution web sites
7 and other processing procedures.

8 A particular data source may contain financial account information, such as
9 financial securities, associated with one or more customers of the corresponding
10 financial institution. Each data element retrieved is associated with a particular
11 identifier, such as an asset identifier or a transaction identifier. An identifier is any
12 number or series of characters assigned to a data element. In a particular
13 embodiment, an identifier is a unique number or series of characters that uniquely
14 and consistently identifies a financial security or similar item. For example, a
15 particular identifier may be associated with a particular holding in an account. In
16 other embodiments, an identifier includes a ticker symbol, a name of a security, or
17 similar information. Similar identifiers are used for data retrieved from multiple
18 financial institutions and multiple financial accounts, thereby allowing the
19 retrieved data to be normalized across the multiple institutions and accounts.
20 When assigning identifiers to data elements, one or more rules may be applied to
21 properly identify the data elements. The particular rules applied to a particular
22 data element may vary depending on the source of the data element.

23 As used herein, the term “data element” refers to any data associated with a
24 financial security (or other item) from any data source. Example data elements
25 include ticker symbols, security names, number of shares, date purchased, date

1 sold, coupon rate, maturity date, security type, industry classification, and the like.
2 As used herein, the terms “account holder”, “customer”, “user”, and “client” are
3 interchangeable. A data element may also refer to a particular account holding,
4 such as a particular stock or a particular bond. “Account holder” refers to any
5 person having access to an account. Various financial account and financial
6 institution examples are provided herein for purposes of explanation. However, it
7 will be appreciated that the systems and procedures described herein can be used
8 with any type of data from any data source. Example financial accounts include
9 savings accounts, money market accounts, checking accounts (both interest-
10 bearing and non-interest-bearing), brokerage accounts, credit card accounts,
11 mortgage accounts, home equity loan accounts, overdraft protection accounts,
12 margin accounts, personal loan accounts, and the like. Example financial
13 institutions include banks, savings and loans, credit unions, mortgage companies,
14 mutual fund companies, lending companies, and stock brokers.

15 Additionally, a data aggregation system may aggregate data from multiple
16 sources, such as multiple financial accounts, multiple email accounts, multiple
17 online award (or reward) accounts, multiple news headlines, and the like.
18 Similarly, the data retrieval and data processing systems and methods discussed
19 herein may be applied to collect data from any type of account containing any type
20 of data. Thus, the methods and systems described herein can be applied to a data
21 aggregation system or any other account management system, and are not limited
22 to the financial analysis systems and procedures discussed in the examples
23 provided herein.

24 Fig. 1 illustrates an example network environment 100 in which various
25 servers, computing devices, and a financial analysis system exchange data across a

1 network, such as the Internet. The network environment of Fig. 1 includes
2 multiple financial institution servers 102 and 106 coupled to a data communication
3 network 108, such as the Internet. Data communication network 108 may be any
4 type of data communication network using any network topology and any
5 communication protocol. Further, network 108 may include one or more sub-
6 networks (not shown) which are interconnected with one another.

7 Another server 104, a client computer 110 and a financial analysis system
8 112 are also coupled to network 108. Financial analysis system 112 is coupled to
9 an asset ID database 116. Asset ID database 116 may also be referred to as an
10 “asset master” or a “security master”. An asset ID is a unique identifier (such as a
11 number or a series of alphanumeric characters) within an identification
12 architecture that is associated with a particular security or a particular class of
13 securities. For example, an asset ID may be associated with a particular stock or a
14 particular bond. An example of an asset ID is a CUSIP (Committee on Uniform
15 Securities Identification Procedures) number. CUSIP is a committee that supplies
16 a unique nine character identification, referred to as a CUSIP number, for each
17 class of security approved for trading in the United States to facilitate clearing and
18 settlement of transactions. Other types of asset IDs include ticker symbols and
19 proprietary identifiers developed by particular financial institutions.

20 Financial analysis system 112 also includes a database 114 that stores
21 various data collected and generated by the financial analysis system. Database
22 114 may also store various identifiers (e.g., ticker symbols), transaction
23 information, and the like. Financial analysis system 112 performs various account
24 analysis functions, data analysis functions, and aggregation functions, as discussed
25 in greater detail below. Although not shown in Fig. 1, financial institution servers

1 102 and 106 may include a database that stores asset identifiers and/or transaction
2 identifiers associated with the particular financial institution.

3 Servers 102-106, client computer 110, and financial analysis system 112
4 may be any type of computing device, such as a desktop computer, a laptop
5 computer, a handheld computer, a personal digital assistant (PDA), a cellular
6 phone, a set top box, or a game console. Client computer 110 is capable of
7 communicating with one or more servers 102-106 to access, for example,
8 information about a financial institution and various user accounts that have been
9 established at the financial institution.

10 The communication links shown between network 108 and the various
11 devices (102, 104, 106, 110, and 112) shown in Fig. 1 can use any type of
12 communication medium and any communication protocol. For example, any of
13 the communication links shown in Fig. 1 may be a wireless link (e.g., a radio
14 frequency (RF) link or a microwave link) or a wired link accessed via a public
15 telephone system or another communication network.

16 Fig. 2 is a block diagram showing example components and modules of
17 financial analysis system 112. A communication interface 202 allows the financial
18 analysis system 112 to communicate with other devices, such as one or more
19 servers or computing devices. In one embodiment, communication interface 202
20 is a network interface to a local area network (LAN), which is coupled to another
21 data communication network, such as the Internet.

22 A database control module 204 allows financial analysis system 112 to
23 store data to database 114 and retrieve data from the database. Financial analysis
24 system 112 also stores various financial institution data 206, which may be used to
25 locate and communicate with various financial institution servers. Financial

1 institution data 206 includes, for example, account balance information,
2 transaction descriptions, transaction amounts, security holdings, asset identifiers
3 and transaction identifiers.

4 A variety of data harvesting scripts 208 are also maintained by financial
5 analysis system 112. For example, a separate data harvesting script 208 may be
6 maintained for each financial institution or other data source from which data is
7 extracted. Data harvesting (also referred to as “screen scraping”) is a process that
8 allows, for example, an automated script to retrieve data from one or more web
9 pages associated with a web site. Data harvesting may also include retrieving data
10 from a data source using any data acquisition or data retrieval procedure.

11 Financial analysis system 112 also includes a data capture module 210 and
12 a data extraction module 212. Data capture module 210 captures data (such as
13 web pages or OFX (Open Financial Exchange) data) from one or more data
14 sources. Data extraction module 212 retrieves (or extracts) data from captured
15 web pages or other data sources. Data extraction module 212 may use one or
16 more data harvesting scripts 208 to retrieve data from a web page.

17 Data capture module 210 may also retrieve data from data sources other
18 than web pages. For example, data capture module 210 can retrieve data from a
19 source that supports the OFX specification or the Quicken Interchange Format
20 (QIF). OFX is a specification for the electronic exchange of financial data
21 between financial institutions, businesses and consumers via the Internet. OFX
22 supports a wide range of financial activities including consumer and business
23 banking, consumer and business bill payment, bill presentment, and investment
24 tracking, including stocks, bonds, mutual funds, and 401(k) account details. QIF
25 is a specially formatted text file that allows a user to transfer Quicken transactions

1 from one Quicken account register into another Quicken account register or to
2 transfer Quicken transactions to or from another application that supports the QIF
3 format.

4 An identification engine 214 analyzes data and various rules to associate
5 identifiers with data or data elements. For example, identification engine 214 can
6 analyze financial account data retrieved from one or more financial institutions.
7 The retrieved data may be obtained by harvesting information from a web site or
8 other data source. Identification engine 214 identifies data elements contained in
9 the financial account data and associates an asset identifier or a transaction
10 identifier with each data element. If an identifier cannot be determined for a
11 particular data element, an exception handling module 216 allows an
12 administrator, developer, or other user to associate an identifier with the particular
13 data element or modify the logic rules associated with the identification.
14 Similarly, if multiple identifiers are determined for a particular data element,
15 exception handling module 216 allows a user to associate a single identifier with
16 the particular data element or modify the logic rules associated with the
17 identification. Exception handling module 216 may also be referred to as an
18 “exception handling tool”. For example, exception handling module 216 allows
19 the user to add new rules, delete rules, or modify rules such that the particular data
20 element will be processed automatically (i.e., without user intervention) in the
21 future by identification engine 214. By continually adding, deleting and
22 modifying rules, the overall performance of the rules in associating identifiers with
23 data elements improves over time.

24 Financial analysis system 112 also includes rules data 218. For example,
25 this rules data is used by identification engine 214 to identify asset identifiers

1 associated with one or more data elements. Rules data 218 may include generic
2 rules and/or one or more sets of rules related to particular financial institutions or
3 other organizations.

4 Although a single identification engine 214 is shown in Fig. 2, alternate
5 embodiments of financial analysis system 112 may include multiple identification
6 engines, such as an asset identification engine, a transaction identification engine,
7 and a proprietary identification engine (e.g., proprietary to a particular financial
8 institution).

9 In particular embodiments, one or more of the components shown in Fig. 2
10 may be omitted from financial analysis system 112, or one or more additional
11 components may be added to financial analysis system 112. Additionally, any of
12 the components shown in Fig. 2 may be combined into another component. For
13 example, data capture module 210 and data extraction module 212 may be
14 combined in a single component. The components shown in Fig. 2 can be
15 implemented in hardware, software, or combinations of hardware and software.

16 Figs. 3A and 3B illustrate a flow diagram of a procedure 300 for retrieving
17 data and associating identifiers with the retrieved data. Initially, the procedure
18 retrieves data from a data source, such as a financial institution (block 302). The
19 procedure then identifies various data elements contained in the retrieved data
20 (block 304). These data elements include, for example, one or more account
21 holdings, one or more account transactions, and other data (or portions of data)
22 retrieved from the data source. The procedure then identifies one or more rules for
23 associating data elements with an identifier (block 306). For example, different
24 sets of rules may be used depending on the data source (or data sources) from
25 which the data elements were retrieved. Since different data sources may use

1 different identifiers and other information to identify data elements, different rules
2 may be necessary to properly identify data elements from the different data
3 sources. For example, different ticker symbols or different naming formats may
4 be used by different data sources. Further, rules may be applied in different orders
5 for different data sources to increase the likelihood of properly identifying the data
6 element and/or to reduce the time required to identify the data element. In certain
7 embodiments, the same set of rules may be associated with two or more different
8 data sources. This identification of rules may be performed by identification
9 engine 214 (Fig. 2).

10 Next, the procedure attempts to associate one or more data elements with an
11 identifier using the rules identified above (block 308). This association may be
12 performed, for example, by identification engine 214 (Fig. 2). As discussed in
13 greater detail below, any number of rules or other information is useful in
14 associating identifiers with data elements. Identifiers include asset identifiers,
15 transaction identifiers, and the like.

16 Procedure 300 continues by determining whether any data elements do not
17 have an associated identifier after processing the retrieved data (block 310). If so,
18 an exception handling module (e.g., module 216 in Fig. 2) is activated to associate
19 identifiers with data elements that do not have associated identifiers (block 312).
20 Additionally, one or more rules may be added or existing rules may be modified to
21 increase the likelihood of successfully associating an identifier with the data
22 elements in the future. Next, the procedure determines whether any data elements
23 have multiple associated identifiers (block 314). This situation occurs when the
24 applied rules indicate two or more possible identifiers that may be associated with
25 a data element. If this occurs, the exception handling module is activated to

1 associate a single identifier with each of the data elements having multiple
2 associated identifiers (block 316). Additionally, one or more rules may be added
3 or existing rules may be modified to increase the likelihood of successfully
4 associating a single identifier with the data elements in the future.

5 After ensuring that one or more data elements have associated identifiers,
6 procedure 300 stores the data elements and the identifiers associated with the data
7 element (block 318). The procedure continues by optionally retrieving additional
8 information regarding the data elements using the associated identifiers (block
9 320). For example, a group of data elements may be associated with a particular
10 asset identifier (also referred to as an “asset code” or an “asset ID”). Additional
11 information regarding this asset may be retrieved from a database or another data
12 source. For example, the procedure may access an asset ID database to obtain
13 more information regarding the particular asset ID. This additional information
14 includes, for example, pricing feeds, industry codes, security size, security type,
15 and the like. This additional information may be obtained from any number of
16 different data sources. In a particular embodiment, an identifier is associated with
17 a single data element. In other embodiments, identifiers are associated with
18 multiple data elements, such as a group or set of data elements.

19 Fig. 4 is a flow diagram illustrating a procedure 400 for retrieving data
20 from multiple financial accounts and associating asset identifiers with the retrieved
21 data based on various rules. Initially, data is retrieved from multiple financial
22 accounts (block 402). The procedure then identifies data elements in the retrieved
23 data (block 404). Procedure 400 continues by identifying generic rules for
24 associating asset identifiers with the data elements (block 406). These rules may
25

1 be related to a group of financial institutions, or a particular industry or
2 organization.

3 Procedure 400 then identifies rules associated with a particular financial
4 institution (block 408). Alternatively, the rules may be associated with a group of
5 financial institutions or another organization. The procedure determines an asset
6 identifier associated with each of the data elements by applying the rules (generic
7 and/or associated with a particular financial institution) to the data elements (block
8 410). The data elements and the associated asset identifiers are then stored for
9 future use (block 412). Although the embodiment of Fig. 4 refers to asset
10 identifiers, similar procedures may be used in alternate embodiments to identify
11 transaction identifiers and other types of information. In these alternate
12 embodiments, the same rules may be used to determine other identifiers or
13 different sets of rules may be identified to associate other identifiers with the data
14 elements.

15 As mentioned above, data is retrieved from one or more data sources, such
16 as financial institutions. In one embodiment, data is retrieved by capturing an
17 HTML (HyperText Markup Language) screen from a financial institution web site.
18 For example, the HTML screen may be a web page associated with the financial
19 institution. Data is then extracted from the HTML screen using a data harvesting
20 script. The extracted data can be normalized, which refers to the process of
21 arranging the extracted data into a standard format. The normalized data is then
22 stored in a database (e.g., database 114 in Fig. 1) for future reference.

23 The normalizing of data is useful when collecting data from multiple
24 sources (e.g., multiple financial institutions). Each financial institution may use
25 different identifiers or other terms for the same type of data. For example, one

1 financial institution may use the identifier "ACME.A" while another financial
2 institution uses the identifier "ACME.C.A" for the same security. By normalizing
3 the data elements, data elements can be grouped in a logical manner. Thus,
4 various financial analysis tools and procedures can analyze data across multiple
5 financial institutions or other data sources. For example, all identifiers related to a
6 particular identifier are normalized to that common identifier. For example, if the
7 identifier is "ACME.A", the related identifier "ACME.C.A" is normalized to the
8 "ACME.A" identifier. This normalization enhances the handling of data from
9 multiple data sources by relating different identifiers associated with the same
10 security to a common identifier.

11 Normalization can be performed by converting an identifier from one
12 format to another (e.g., converting "ACME.C.A" to "ACME.A"). Alternatively,
13 one or more rules may associate different holdings or ticker symbols with the same
14 asset identifier. For example, a first rule may associate "ACME.C.A" with asset
15 identifier "12345". Similarly, a second rule may associate "ACME.A" with the
16 same asset identifier "12345".

17 As mentioned above, data harvesting (or screen scraping) is a process that
18 allows a script to retrieve data from a web site and store the retrieved data in a
19 database. Data harvesting scripts are capable of navigating web sites and
20 capturing individual HTML pages. For example, JavaScript and images may be
21 removed from the HTML pages or converted into HTML text if it contains account
22 information. A parser then converts the HTML data into a field-delimited XML
23 format. The XML data communicates with enterprise java beans (EJBs) through
24 an XML converter. EJBs perform a series of SQL queries that populate the data
25 into the database.

1 When retrieving data from a data source other than an HTML screen, the
2 data source may communicate data using the OFX standard, the QIF format, or
3 any other data format. Data is retrieved from the source and a procedure identifies
4 data of interest. The data of interest may be, for example, data associated with a
5 particular financial institution. The identified data is then normalized and stored in
6 a database. The database may contain data related to other customers and/or data
7 collected from other sources (such as HTML screens).

8 One or more sets of rules (also referred to as “search patterns”) may be
9 applied when determining identifiers associated with a data element. Different
10 sets of rules may be associated with different financial institutions or with
11 different types of data elements. In a particular embodiment, a first set of rules
12 includes generic rules that may be applied to different types of data elements
13 associated with different financial institutions. In this embodiment, other sets of
14 rules are specific to a particular financial institution or to a particular type of data
15 element. In other embodiments, any number of rules (or sets of rules) may be
16 used when determining identifiers associated with data elements.

17 Fig. 5 is a flow diagram illustrating a procedure 500 for applying various
18 rules or search patterns to determine an identifier associated with a data element.
19 Initially, procedure 500 identifies a first generic rule (block 502) from a set of one
20 or more generic rules. The procedure applies the selected generic rule to the
21 retrieved data element (block 504). Next, the procedure determines whether
22 application of the selected generic rule has resulted in a single identifier being
23 matched with (or associated with) the retrieved data element (block 506). If so,
24 the identifier is associated with the data element and the procedure is complete for
25 that particular data element (block 508).

1 If a single identifier match has not occurred in block 506, the procedure
2 determines whether there are additional generic rules to apply (block 510). If so,
3 the procedure identifies the next generic rule (block 512) and returns to block 504
4 to apply the next generic rule to the received data element. If all generic rules
5 have been applied, the procedure continues from block 510 to block 514, which
6 identifies a first financial institution-specific (FI-specific) rule. FI-specific rules
7 are associated with a particular financial institution and incorporate information
8 specific to the financial institution, such as security naming conventions, ticker
9 symbol formats, and the like. For example, a particular FI-specific rule may
10 change the abbreviation “FD” to “FUND” to provide a consistent naming
11 convention among multiple data sources.

12 Procedure 500 continues by applying the selected FI-specific rule to the
13 retrieved data element (block 516). The procedure then determines whether
14 application of the selected FI-specific rule has resulted in a single identifier being
15 matched with (or associated with) the retrieved data element (block 518). If so,
16 the identifier is associated with the data element and the procedure is complete for
17 that particular data element (block 508).

18 If a single identifier match has not occurred in block 518, the procedure
19 determines whether there are additional FI-specific rules to apply (block 520). If
20 so, the procedure identifies the next FI-specific rule (block 512) and returns to
21 block 516 to apply the next FI-specific rule to the received data element. If all FI-
22 specific rules have been applied, the procedure continues from block 520 to block
23 524, which generates an indication that a single match was not identified by
24 applying the various generic and FI-specific rules. Although the example of Fig. 5
25 applies generic rules before FI-specific rules, alternate embodiments may apply

1 rules in any order. For example, one or more FI-specific rules may be applied
2 before applying one or more generic rules. In other embodiments, one or more FI-
3 specific rules are applied instead of any generic rules.

4 In a particular implementation, application of each rule may narrow a pool
5 of possible identifiers that may be associated with a particular data element. For
6 example, application of a first rule may narrow a pool of possibilities to ten
7 possible identifiers. The second rule is then applied to these ten possible
8 identifiers, which narrows the pool to three possible identifiers. The third rule is
9 applied to those three possible identifiers, but may not further reduce the size of
10 the pool. Finally, a fourth rule is applied to the three possible identifiers and
11 results in a single identifier that is associated with the data element. In other
12 examples, any number of rules may be applied before a single identifier is
13 determined.

14 In another implementation, each rule is applied to the entire universe of
15 possible identifiers. Thus, if the first rule does not identify a single identifier, the
16 next rule is applied. Each subsequent rule is more specific or combines one or
17 more selection features of the previous rules. These rules may be prioritized to
18 efficiently and accurately identify the proper identifier for one or more data
19 elements.

20 In some situations, application of all rules leaves a pool of two or more
21 possible identifiers. In this situation, a user may manually determine which
22 identifier is the correct identifier for the data element. Additionally, a new rule
23 may be developed or an existing rule may be modified to handle this situation in
24 the future.

25

Fig. 6 illustrates an example set of rules 600 used to associate data elements with identifiers. A first column 602 identifies a ranking or priority associated with each rule identified in a second column 604. In the example of Fig. 6, a first rule converts ticker symbols to a standard format. For example, if a particular financial institution represents ticker symbols in a particular format, the format of ticker symbols associated with that financial institution is converted into a standard format used for all ticker symbols from any data source. Thus, if the data element contains a non-standard ticker symbol, that ticker symbol is converted to a standard format. The next rule attempts to match the data element with a particular ticker symbol from a list of all possible ticker symbols. If a single match is not identified, the next rule converts non-standard names in the data element to a standard format. The next rule attempts to match the data element with a name from a list of possible security names. If an exact match is not identified, the next rule determines whether a match of at least three words in the name is found. If not, the next rule attempts to match the exact description with a description from a list of possible security descriptions. If there is still no match, the last rule shown attempts to match at least ten words in the description.

In alternate embodiments, a set of rules may include any number of rules. Further, multiple sets of rules may be applied to a particular data element when attempting to associate the data element with an identifier. Further, the order in which rules are applied may vary. For example, in Fig. 6, the second rule (match ticker symbol) may be applied first. If that rule does not identify a single identifier, then the first rule (convert ticker symbols to standard format) is applied to the data element. Any number of rules may be applied in any order when attempting to identify an identifier associated with a data element.

1 Fig. 7 is a block diagram showing pertinent components of a computer 700
2 in accordance with the invention. A computer such as that shown in Fig. 7 can be
3 used, for example, to perform various procedures such as those discussed herein.
4 Computer 700 can also be used to access a data source or other device to access
5 various financial information. The computer shown in Fig. 7 can function as a
6 server, a client computer, or a financial analysis system, of the types discussed
7 herein.

8 Computer 700 includes at least one processor 702 coupled to a bus 704 that
9 couples together various system components. Bus 704 represents one or more of
10 any of several types of bus structures, such as a memory bus or memory controller,
11 a peripheral bus, and a processor or local bus using any of a variety of bus
12 architectures. A random access memory (RAM) 706 and a read only memory
13 (ROM) 708 are coupled to bus 704. Additionally, a network interface 710 and a
14 removable storage device 712, such as a floppy disk or a CD-ROM, are coupled to
15 bus 704. Network interface 710 provides an interface to a data communication
16 network such as a local area network (LAN) or a wide area network (WAN) for
17 exchanging data with other computers and devices. A disk storage 714, such as a
18 hard disk, is coupled to bus 704 and provides for the non-volatile storage of data
19 (e.g., computer-readable instructions, data structures, program modules and other
20 data used by computer 700). Although computer 700 illustrates a removable
21 storage 712 and a disk storage 714, it will be appreciated that other types of
22 computer-readable media which can store data that is accessible by a computer,
23 such as magnetic cassettes, flash memory cards, digital video disks, and the like,
24 may also be used in the example computer.

1 Various peripheral interfaces 716 are coupled to bus 704 and provide an
2 interface between the computer 700 and the individual peripheral devices.
3 Example peripheral devices include a display device 718, a keyboard 720, a mouse
4 722, a modem 724, and a printer 726. Modem 724 can be used to access other
5 computer systems and devices directly or by connecting to a data communication
6 network such as the Internet.

7 A variety of program modules can be stored on the disk storage 714,
8 removable storage 712, RAM 706, or ROM 708, including an operating system,
9 one or more application programs, and other program modules and program data.
10 A user can enter commands and other information into computer 700 using the
11 keyboard 720, mouse 722, or other input devices (not shown). Other input devices
12 may include a microphone, joystick, game pad, scanner, satellite dish, or the like.

13 Computer 700 may operate in a network environment using logical
14 connections to other remote computers. The remote computers may be personal
15 computers, servers, routers, or peer devices. In a networked environment, some or
16 all of the program modules executed by computer 700 may be retrieved from
17 another computing device coupled to the network.

18 Typically, the computer 700 is programmed using instructions stored at
19 different times in the various computer-readable media of the computer. Programs
20 and operating systems are often distributed, for example, on floppy disks or CD-
21 ROMs. The programs are installed from the distribution media into a storage
22 device within the computer 700. When a program is executed, the program is at
23 least partially loaded into the computer's primary electronic memory. As
24 described herein, the invention includes these and other types of computer-
25 readable media when the media contains instructions or programs for

1 implementing the steps described below in conjunction with a processor. The
2 invention also includes the computer itself when programmed according to the
3 procedures and techniques described herein.

4 For purposes of illustration, programs and other executable program
5 components are illustrated herein as discrete blocks, although it is understood that
6 such programs and components reside at various times in different storage
7 components of the computer, and are executed by the computer's processor.
8 Alternatively, the systems and procedures described herein can be implemented in
9 hardware or a combination of hardware, software, and/or firmware. For example,
10 one or more application specific integrated circuits (ASICs) can be programmed to
11 carry out the systems and procedures described herein.

12 Although the description above uses language that is specific to structural
13 features and/or methodological acts, it is to be understood that the invention
14 defined in the appended claims is not limited to the specific features or acts
15 described. Rather, the specific features and acts are disclosed as example forms of
16 implementing the invention.